

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3586777号

(P3586777)

(45) 発行日 平成16年11月10日(2004.11.10)

(24) 登録日 平成16年8月20日(2004.8.20)

(51) Int. Cl.<sup>7</sup>

F I

G06F 3/16  
G10L 15/24  
G10L 15/28

G06F 3/16 320B  
G10L 3/00 571G  
G10L 3/00 571C

請求項の数 7 (全 14 頁)

(21) 出願番号 特願平8-193418  
(22) 出願日 平成6年8月17日(1994.8.17)  
(65) 公開番号 特願平8-63330  
(43) 公開日 平成8年3月8日(1996.3.8)  
審査請求日 平成12年7月3日(2000.7.3)

(73) 特許権者 000005223  
富士通株式会社  
神奈川県川崎市中原区上小田中4丁目1番  
1号  
(74) 代理人 100078868  
弁理士 河野 登夫  
(72) 発明者 廣尾 信之  
神奈川県川崎市中原区上小田中1015番  
地 富士通株式会社内

審査官 逆藤 尊志

(56) 参考文献 特開昭61-275972(JP,A)  
特開平06-110461(JP,A)

最終頁に続く

(54) 【発明の名称】 音声入力装置

## (57) 【特許請求の範囲】

## 【請求項1】

音声入力部と、音声以外の情報を入力する入力装置と、前記音声入力部から入力された音声情報に所定の処理を施す音声処理部とを備えた音声入力装置において、  
前記入力装置の信号に基づいて前記入力装置が使用中であるか否かを検出し、使用中であると検出された入力装置に応じて、入力された音声情報に施すべき処理の内容を判定し、判定した結果に応じて前記音声処理部を制御する判定部とを具備することを特徴とする音声入力装置。

## 【請求項2】

前記判定部は、使用中であると検出された入力装置に入力された音声情報の内容が設定された設定ファイルを有し、

使用中であると検出された入力装置と前記設定ファイルに基づいて、入力された音声情報に施すべき処理の内容を判定し、判定した結果に応じて前記音声処理部を制御することを特徴とする請求項1記載の音声入力装置。

## 【請求項3】

前記設定ファイルは、音声情報をデフォルト又は使用中のアプリケーションと対応付けて設定してあり、

前記判定部は、使用中のアプリケーションの信号を検知し、使用中のアプリケーションの有無とアプリケーションが使用中であるか否かを判定し、

使用中であると検出された入力装置、使用中であると判定されたアプリケーション、及

19

20

A1-

び前記設定ファイルに基づいて、入力された音声情報に施すべき処理の内容を判定し、判定した結果に応じて前記音声処理部を制御することを特徴とする請求項2に記載の音声入力装置。

【請求項4】

前記音声処理部は複数の音声辞書と、この複数の音声辞書のうちのいずれか一つ又は複数を選択する辞書切替部とを備え、

前記判定部は、判定した結果に応じて前記辞書切替部に音声辞書の選択指令を出力し、前記辞書切替部は、前記選択指令に基づいて音声辞書を切り替えることを特徴とする請求項1乃至3のいずれか一項に記載の音声入力装置。

【請求項5】

音声入力部と、音声以外の情報を入力する入力装置と、前記音声入力部から入力された音声情報に所定の処理を施す音声処理部と、該音声処理部で処理された結果を出力する出力部とを備えた音声入力装置において、

前記入力装置の信号に基づいて前記入力装置が使用中であるか否かを検出し、使用中であると検出された入力装置に応じて、入力された音声情報に施すべき処理内容及びこの処理結果の出力態様を判定し、この判定結果に応じて前記音声処理部及び前記出力部を制御する判定部とを具備することを特徴とする音声入力装置。

【請求項6】

入力された音声情報に対する音声処理部の認識結果が予め定めた単語である場合に、前記出力部は判定部の判定結果の如何にかかわらず、予め定めた態様の出力を行うべく動作するようにしてあることを特徴とする請求項5に記載の音声入力装置。

【請求項7】

前記入力装置はキーボード及び／又はマウスであることを特徴とする請求項1乃至6のいずれか一項に記載の音声入力装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】

本発明は入力された音声情報に施すべき処理内容の変更、また入力された音声情報の出力内容の変更をスイッチ等を用いずに自動的に切換え可能とした音声入力装置に関する。

【0002】

【従来の技術】

図7は従来における音声入力装置の構成を示すブロック図であり、図中1はマイク等の音声入力部、2a、2b…2nはキーボード、マウス等音声以外の他の情報を入力する入力装置を示している。

音声入力部1から入力された音声情報は音声認識部5へ入力される。

音声認識部5は、予めスイッチ21にて入力される音声情報、例えばテキスト情報、コマンド情報等々に応じた処理モードに設定されており、処理モードがテキスト情報処理モードである場合には辞書格納部22からテキスト情報処理用辞書を読み出し、これに基づいて、またコマンド情報処理モードである場合には辞書格納部22からコマンド情報処理用辞書を読み出し、これに基づき入力された音声情報の認識処理を行い、認識結果を処理結果出力部6へ出力する。

処理結果出力部6も予めスイッチ21にて入力される音声情報に対応した出力モードに設定されており、入力された認識結果を、例えばテキストとして、又はコマンドとして夫々他の入力装置2a～2nからの入力情報と共に出力する。

【0003】

【発明が解決しようとする課題】

ところで、音声入力部1を通じて入力されてくる対象は、例えば文章等の文字情報である場合、又アプリケーション、ウィンドウマネージャ、OSに対する操作命令である場合、又は音声波形データである場合等その時々によって変化する。

【0004】

19

20

30

40

50

このような種々の入力対象に対し音声認識部5において施すべき処理の内容、処理の手順も自づと異なるから、音声認識部5を夫々の入力対象に適応した処理モードに切換える必要があり、従来にあつては、スイッチ21を手動、又は音声入力により切換えて処理モードの設定を行っていた。この点は処理結果出力部6においても同様である。

【0005】

しかしスイッチ21を、例えば手動により切換えるには使用者は使用中のキーボード、又はマウス等から一旦手を離さざるを得ず、キーボード、マウスの操作が中断されることとなり、また音声入力により切換えるには、当然切換えのための特別なコマンドを登録しておく必要がある上、ノイズ、その他入力音声以外の周辺での会話等に起因する誤認識が生じ、操作者が期待していない時点で突発的に処理モード、出力モードの切換えが行われることがある等の不都合があつた。

【0006】

本発明はかかる事情に鑑みなされたものであつて、その目的とするところは入力音声に施すべき処理の変更、及び出力態様の変更を操作者に特別の操作を要求することなく自動的にに行い得るようにすることにある。

本発明の他の目的は音声処理部において入力された音声情報に対して音声認識処理を行い場合にテキスト、コマンド等入力された音声情報に応じて音声辞書の切換えを自動的にに行い得るようにすることにある。

【0007】

本発明の更に他の目的は、入力された音声情報に対して、音声処理を施すことなく出力する場合にも判定部にて、これを自動的に判定して出力部に対する制御を可能とすることにある。

本発明の更に他の目的は、キーボード、マウス等通常のコンピュータに備えられているものの使用状況及び／又は使用履歴に基づいて判定部が判定を行うこととすることで広範囲にわたる適用を可能とすることにある。

【0008】

本発明の更に他の目的は、入力された音声情報が予め定めた単語である場合には判定部の判定結果の如何にかかわらず、予め定めた態様の出力を出力部から行わせることで、処理の効率化を図ることにある。

本発明の更に他の目的は、判定部における判定処理の内容を必要に応じて変更可能とすることで、適用範囲を更に拡大可能ならしめることにある。

【0009】

【課題を解決するための手段】

以下本発明の原理を図1に示す原理図に基づき説明する。

図1は本発明に係る音声入力装置の原理図であり、図中1は音声入力部、2a、2b～2nはキーボード、マウス等、音声以外の情報を入力する入力装置を示している。

音声入力部1から入力された音声情報はデジタル情報として音声処理部5へ入力される。

【0010】

一方入力装置2a～2nの使用状況及び／又は使用履歴が逐次判定部7へ取り込まれており、判定部7はこれら使用状況及び／又は使用履歴に基づいて、予め設定された判定処理の内容、即ちアルゴリズムに従って音声入力部1から現に入力されつつある音声情報又は次に入力される音声情報が如何なる内容のものか、例えばテキスト入力、又はコマンド入力か、又は音声処理部において何ら処理を施す必要のないデータか等を判定し、この判定結果に基づいて音声処理部5及び出力部6へ夫々所定の指令を与える。

【0011】

一般に、例えばキーボードの入力に熟練した操作者の場合、音声入力により文章等を入力するよりも、キーボードを使用して入力の方が処理を迅速に行えるのが普通である。従つてキーボードを使用している際、熟練した操作者においてあえて音声入力したいと考えるような対象は、例えばウィンドウのオープン、アプリケーションのモード変更等の操作

命令であることが多い。一方マウスを使用中の場合には、文章等を入力するには一旦マウスから手を離し、キーボードを使用して文章を入力し、再びマウスに手を戻す動作が必要となることから、操作命令に限らず音声入力により文字情報の入力を行いたいと欲する場合が多い。

つまり音声以外の情報を入力する入力装置であるキーボード、マウスの使用履歴、使用状況を把握することで、入力音声に対して音声処理部5で施すべき処理内容、出力部6の出力態様を判断することが可能となるのである。

#### 【0012】

音声処理部5に対しては、入力される音声情報に対し、音声認識処理を施すべきか否か、また音声認識処理を施すべき場合にはテキストとして、又はコマンドとして認識処理を行うべきか否かの指令を与え、音声処理部5を制御する。

また出力部6に対しては、音声処理部5から与えられる認識結果がテキストである場合にはテキストとして出力すべく、又はコマンドである場合はコマンドとして出力すべく、更に音声処理部5において何ら処理を施されなかった内容については、例えばこれを波形エディタへ出力すべく指令を与え、出力部6を制御する。

#### 【0013】

これによって音声処理部5は判定部7からの指令に従って入力された音声情報に対応可能にモード設定され、入力された音声情報に所定の音声処理を施して、又は処理を施すことなくこれを出力部6へ出力する。

また出力部6は同じく判定部7からの指令に従って音声処理部5からの入力が、例えばテキスト入力の場合にはテキストとして、またコマンド入力の場合にはコマンド入力として、他の入力装置2a、2b…2nからの入力と同様、ワードプロセッサ、波形エディタ等へ出力する。

#### 【0014】

なお、入力された音声情報の認識結果が予め定めた特定の単語等である場合は出力部は予め定めた態様の出力を他の態様に優先して行うこととしてもよい。

また、操作者は判定部7の判定処理内容は任意に変更可能であって操作者は判定結果を適用対象に応じて変更させることで適用可能範囲を拡大し得るようにしてある。

#### 【0015】

第1の発明に係る音声入力装置は、音声入力部と、音声以外の情報を入力する入力装置と、前記音声入力部から入力された音声情報に所定の処理を施す音声処理部とを備えた音声入力装置において、前記入力装置の信号に基づいて前記入力装置が使用中であるか否かを検出し、使用中であると検出された入力装置に応じて、入力された音声情報に施すべき処理の内容を判定し、判定した結果に応じて前記音声処理部を制御する判定部とを具備することを特徴とする。

#### 【0016】

第2の発明に係る音声入力装置は、第1の発明において、前記判定部は、使用中であると検出された入力装置に入力された音声情報の内容が設定された設定ファイルを有し、使用中であると検出された入力装置と前記設定ファイルに基づいて、入力された音声情報に施すべき処理の内容を判定し、判定した結果に応じて前記音声処理部を制御することを特徴とする。

#### 【0017】

第3の発明に係る音声入力装置は、第2の発明において、前記設定ファイルは、音声情報をデフォルト又は使用中のアプリケーションと対応付けて設定してあり、前記判定部は、使用中のアプリケーションの信号を検知し、使用中のアプリケーションの有無とアプリケーションが使用中であるか否かを判定し、使用中であると検出された入力装置、使用中であると判定されたアプリケーション、及び前記設定ファイルに基づいて、入力された音声情報に施すべき処理の内容を判定し、判定した結果に応じて前記音声処理部を制御することを特徴とする。

#### 【0018】

第4の発明に係る音声入力装置は、第1乃至第3の発明において、前記音声処理部は複数の音声辞書と、この複数の音声辞書のうちのいずれか一つ又は複数を選択する辞書切替部とを備え、前記判定部は、判定した結果に応じて前記辞書切替部に音声辞書の選択指令を出力し、前記辞書切替部は、前記選択指令に基づいて音声辞書を切り替えることを特徴とする。

【0019】

第5の発明に係る音声入力装置は、音声入力部と、音声以外の情報を入力する入力装置と、前記音声入力部から入力された音声情報に所定の処理を施す音声処理部と、該音声処理部で処理された結果を出力する出力部とを備えた音声入力装置において、前記入力装置の信号に基づいて前記入力装置が使用中であるか否かを検出し、使用中であると検出された入力装置に応じて、入力された音声情報に施すべき処理内容及びこの処理結果の出力態様を判定し、この判定結果に応じて前記音声処理部及び前記出力部を制御する判定部とを具備することを特徴とする。

19

【0020】

第6の発明に係る音声入力装置は、第5の発明において、入力された音声情報に対する音声処理部の認識結果が予め定めた単語である場合に、前記出力部は判定部の判定結果の如何にかかわらず、予め定めた態様の出力を行うべく動作するようにしてあることを特徴とする。

【0021】

第7の発明に係る音声入力装置は、第1乃至第6の発明において、前記入力装置はキーボード及び／又はマウスであることを特徴とする。

20

【0022】

【作用】

第1の発明にあっては、音声以外の情報を入力する入力装置が使用中であるか否かに応じて判定部が入力音声に施すべき処理を自動的に判定して音声処理部に対して指示することとなり、操作者は処理内容の指示を必要としない。

【0023】

第2の発明にあっては、音声以外の情報を入力する入力装置が使用中であるか否か、及び設定ファイルの内容に基づいて判定部が入力音声に施すべき処理を自動的に判定して音声処理部に対して指示することとなり、操作者は処理内容の指示を必要としない。

30

【0024】

第3の発明にあっては、音声以外の情報を入力する入力装置が使用中であるか否か、及び設定ファイルの内容に基づいて判定部が入力音声に施すべき処理を自動的に判定して音声処理部に対して指示することとなり、操作者は処理内容の指示を必要としない。

【0025】

第4の発明にあっては、音声処理部において音声認識のために用いる複数の辞書を辞書切替部にて自動的に切替え可能となる。

【0026】

第5の発明にあっては、音声以外の情報を入力する入力装置が使用中であるか否かに基づいて判定部が判定結果に応じて音声処理部、出力部夫々に対し、指令を出力することでこれらに対する制御を自動的に行うことが可能となる。

40

【0027】

第6の発明にあっては、判定部の判定結果の如何にかかわらず、予め定めた認識結果に対し、出力部に予め定めた態様の出力を行わせることで、誤動作を低減すると共に、操作性を向上し得る。

【0028】

第7の発明にあっては、キーボード、マウスが使用中であるか否かに関する情報を用いることで、キーボード、マウスを備える汎用コンピュータへの適用が可能となる。

【0029】

【実施例】

50

(実施例1)

以下本発明をその実施例を示す図面に基づき具体的に説明する。

図2は本発明に係る音声入力装置を図形編集機能付のワードプロセッサ11に適用した場合の構成を示すブロック図であり、図中1はマイク等にて構成された音声入力部、2a、2bは音声以外の情報を入力するキーボード、マウス等の入力装置を示している。

音声入力部1より入力された音声情報はA/D変換部3でアナログ信号をデジタル信号に変換されて、音声認識部として構成された音声処理部5へ入力される。

【0030】

一方音声以外の情報を入力する入力装置2a、2bからの入力情報はワードプロセッサ11へ入力される他、逐次判定部7へ取り込まれる。

判定部7はキーボード、マウス等の入力装置2a、2bからの信号に基づき予め設定した判定処理内容、即ちアルゴリズムに従いこれらの使用状況及び／又は使用履歴を認識し、音声入力部1を通じて現に入力され、また後に入力されてくる音声情報の内容及び入力される音声情報に対して施すべき処理の内容を判定する。

具体的には入力されてきた音声情報がテキスト情報か、コマンド情報か、並びに夫々の情報に対し音声処理部5で施すべき処理の内容及び出力部6からの出力様子を判定し、夫々に応じた指令を辞書切替部8及び出力部6へ与える。

【0031】

なお、キーボード、マウス等の各入力装置2a、2bにその使用の有無を検出するセンサが付設されている場合、このセンサ出力を判定部7に取り込み、これらの使用状況、使用履歴を認識し、判定を行うこととしてもよい。

辞書切替部8は判定部7からの指令によりテキスト用辞書、又はコマンド用辞書10を音声処理部5へ読み出す。

【0032】

音声処理部5は前記判定部7からの指令に基づき動作される辞書切替部8にて選択的に切替えられたテキスト用辞書9又は／コマンド用辞書10を読み出し、これらに基づいて、音声情報の認識処理を行い、認識結果を出力部6へ出力する。出力部6は前記判定部7からの指示に基づき音声情報がテキスト入力の場合にはテキストとして、またコマンド入力の場合にはコマンドとしてこれをワードプロセッサ11へ出力する。

【0033】

次に本発明装置の動作を図3に示すフローチャートに従って説明する。

図3は判定部7が現在使用中の入力装置が何であるかに基づいて判定を行う場合の処理過程を示すフローチャートであり、先ず使用中の入力装置2a、2bは何れかを判断し（ステップS1）、使用中の入力装置がマウスの場合には入力される音声情報はワードプロセッサ11で編集集中の文書に対するテキスト入力と判定し（ステップS2）、またキーボードである場合には、入力される音声情報はワードプロセッサ11に対するコマンド入力と判定し（ステップS3）、夫々の判定に基づき辞書切替部8及び出力部6へ対応する指示を出力する。

【0034】

次に具体例を挙げて処理内容を説明する。

例えば操作者がキーボードを使用してワードプロセッサ11により文章を作成中である場合、文章のバックアップを採るべく「セーブ」と発声すると、判定部7は操作者がキーボード使用中であることを認識し、入力された音声情報が前述の如くワードプロセッサ11に対するコマンド入力と判定し、辞書切替部8に対しコマンド用辞書10を選択すべく指令を出力し、また出力部6に対しては音声認識部の認識結果をコマンドとして、ワードプロセッサ11へ出力すべく指示する。

【0035】

この結果、音声認識部として構成された音声処理部5においては入力された音声情報を、コマンド用辞書10を用いて「セーブ」と認識し、その認識結果を出力部6へ出力する。出力部6は認識結果「セーブ」をコマンド「save」としてワードプロセッサ11へ出

かし、ワードプロセッサ 11 はコマンド「save」を受けて編集中の文書のセーブを行う。

【0036】

また操作者がワードプロセッサ 11 にて図形編集を行っているものとして、その図形中の所定部分に、例えば「日本語」というテキストを書入れるべく、先ず「日本語」を入れたい位置をマウスにて指定し、「日本語」と発声したとする。判定部 7 は操作者がマウスの使用中であることを認識し、前述した如く入力された音声をワードプロセッサ 11 の編集中の文書に対するテキスト入力と判定し、辞書切替部 8 にテキスト用辞書 9 を選択すべく指示し、また出力部 6 に対してはテキスト表示として出力すべく指示する。

【0037】

これによって音声処理部 5 は入力された音声情報をテキスト用辞書 9 を用いて「日本語」と認識し、この認識結果を出力部 6 へ出力する。出力部 6 は「日本語」をテキストとしてワードプロセッサ 11 へ出力し、ワードプロセッサ 11 はマウスによる指示位置にテキストである「日本語」を挿入表示する。

【0038】

(実施例 2)

実施例 2 は波形エディタ 12 を用いて入力された音声情報に対する編集を行っており、入力された音声情報に対し音声認識部として構成された音声処理部 5 が特別な処理を施す必要のない場合を示している。

図 4 は本発明の実施例 2 の構成を示すブロック図である。この実施例 2 においては A/D 変換部 3 と音声認識部として構成された音声処理部 5 との間に音声記憶部 4 を介装し、判定部 7 からの指示は辞書切替部 8、出力部 6 の他に、この音声記憶部 4 へも出力するようにしてある。また波形エディタ 12 はキーボード、マウス等の入力装置 2 a、2 b 夫々からの出力の他に、出力部 6 からの出力が入力され、波形エディタ 12 からは波形エディタ使用中であることを示す信号が判定部 7 へ与えられるようにしてある。

【0039】

判定部 7 は、キーボード、マウス等の入力装置 2 a、2 b の使用を示す信号と、波形エディタ 12 からの波形エディタの使用を示す信号とに基づき、入力された音声情報の内容が波形編集のためのデータであることを認識し、音声記憶部 4 へ音声記憶すべく指令を出力し、また出力部 6 に対してはその波形を波形エディタ 12 へ出力すべく指令を出力する。

図 5 は判定部 7 の処理過程を示すフローチャートである。先ず、入力された音声情報が音声記憶部 4 に録音中か否かを判定し（ステップ S 11）、録音中であれば入力された音声情報（波形）を出力するのみで、これに対する認識処理を行わない対象であると判定する（ステップ S 12）、一方入力された音声情報を録音していない場合には、使用中の入力装置はキーボードか、又はマウスかを判断する（ステップ S 13）。

【0040】

キーボードの場合には入力された音声情報をコマンド入力と判定し（ステップ S 14）、またマウスを使用中の場合には文字入力の要求があるか否かを判断し（ステップ S 15）、無い場合には入力された音声情報をコマンド入力と判定し（ステップ S 14）、また有る場合には入力された音声情報はテキスト入力と判定する（ステップ S 16）。

【0041】

具体的に操作者が自らの声をマイクを通じて入力（録音）し、その波形を編集し、編集結果をファイルに保存すべく作業中の場合について説明する。

操作者はマイクに向かって発声し、自らの声の録音を開始する。このような状態下では波形エディタ 12 から判定部 7 に対し、音声の録音中である旨の情報が入力される。これによって判定部 7 は音声処理部 5 で入力された音声情報に対し、音声の認識処理を施す必要がなく、単にその波形を出力するのみでよいと判定する。

判定部 7 はこの判定に基づき音声記憶部 4 に対し入力された音声情報を録音すべく指令し、また出力部 6 に対しては入力された音声波形をそのまま波形エディタ 12 へ出力すべく

指示する。なお辞書切替部8に対しては音声認識処理を必要としないことから指令は出力されない。

【0042】

この結果、A/D変換部3にてデジタル化された音声情報は音声記憶部4にて録音された後、直接出力部6へ出力され、また出力部6は入力された音声波形を波形エディタ12へ出力する。

操作者は発声の録音が終了すると波形の編集を開始する。波形エディタ12は操作者が波形の区間をマウスを用いて指定し、「エコー」と発声すると指定された波形に対しエコー処理を施し、また「クリア」と発声したとすると指定された波形を消去すべく処理を行う。

19

【0043】

即ち、現在キーボードの使用中之ある場合、判定部7はキーボードからの使用中であることを示す信号及び波形エディタ12を通じて入力される信号に基づき入力された音声情報はコマンドであると判定する。

これに従って判定部7は音声記憶部4に対し、音声処理部5へ音声を送るべく指令し、また辞書切替部8に対してはコマンド用辞書10を選択すべく指令し、出力部6に対してはコマンドを波形エディタ12へ送るべく指令する。

【0044】

この結果、音声処理部5はコマンド用辞書10を用いて入力された音声情報に対する認識処理を行い、入力音声である、例えば「エコー」又は「クリア」を認識し、これを出力部6へ出力する。

20

出力部6は認識結果である「エコー」又は「クリア」をコマンドとして波形エディタ12へ送り、このコマンドが実行される。

次に操作者が編集した内容を保存すべく「セーブ」と発声したとする。この「セーブ」が名称未設定ファイル、換言すれば新規ファイルである場合、波形エディタ12はファイルの名称を要求する。そこでファイル名として「自分の声」と発声した場合、マウスを使用中であっても波形エディタ12はテキスト入力を要求するから判定部7が入力された音声情報をテキストと判定する。

【0045】

判定部7は辞書切替部8に対しテキスト用辞書9を選択すべく指令を出力し、また出力部6に対してはテキストとしての「自分の声」を出力すべく指示する。この結果、音声処理部5はテキスト用辞書9を用いて音声情報に対する認識処理を行い、これを出力部6へ出力する。出力部6は認識結果である「自分の声」をテキストとして波形エディタ12へ出力し、ファイル名である「自分の声」が波形エディタ12へ入力され、セーブされる。

30

このような実施例2にあつてはファイル名の如き文字入力、又は「エコー」の如きディレイタイムの数値入力等は操作中のマウスからキーボードに手を移さなくても音声入力により入力が可能となる。

【0046】

なお、実施例1、2のいずれの場合について、判定部7の判定結果が如何なるものであつても、音声認識の結果が予め定めた「特定単語」である場合には出力部6は予め定めた所定の出力制御を行うこととしてもよい。

40

例えば特定単語がウィンドウマネージャ、OSに対する操作指令である「リサイズ」又は所定の人名、例えば「田中」である場合、出力部6は「リサイズ」の場合にあつてはウィンドウのサイズ変更のための操作指令をウィンドウマネージャ、OSへ出力する。

【0047】

「リサイズ」の場合、所定のウィンドウのもとでアプリケーションを操作中であつて、判定部7が入力された音声情報をアプリケーションへのコマンドと判定した場合、実質的に入力音声に対する処理内容の優先順位を認識結果を利用して設定しているのと等価となり、操作性が格段に向上する。

また、広く使われている人名である、例えば「田中」が音声入力部1から入力された場合

50



、これを「無視」するように判定部の処理内容を設定することで周囲から「田中」の音声  
が頻繁に混入する虞れがある場合においてもこれによる誤認を避け得ることとなる。

【0048】

(実施例3)

実施例1、2では判定部7に対して入力装置2a、2bの使用状況、使用履歴に基づき如  
何なる判定を行わせるかの判定処理内容は、音声入力システムの始動に際して初期設定さ  
れる場合について説明したが、この実施例3では任意の時点で再設定することが可能とな  
っている。

【0049】

図6(a)は判定部7における判定処理内容、即ちアルゴリズムの初期設定処理過程、図  
6(b)はアルゴリズムの設定変更処理過程夫々のフローチャートである。

先ず、アルゴリズムの初期設定は音声入力装置の起動時に初期設定ファイルが存在するか  
否かを判断し(ステップS21)、存在しない場合は「固有の設定」、例えばキーボ  
ード使用時はコマンド入力と、またマウス使用時はテキスト入力とする判定処理の設定を行  
う(ステップS22)。

【0050】

また存在する場合、換言すればユーザーが好みに応じて設定する設定ファイルが存在する  
場合には前期「固有の設定」に優先して、判定部7は初期設定ファイルを読み込み(ステ  
ップS23)、この初期設定ファイルの内容に従って設定を行い(ステップS24)、  
設定ファイルに現在の設定内容を保存する(ステップS25)。

【0051】

一方再設定を行う場合には設定ファイルをユーザーがエディタ等を用いて変更し(ステッ  
プS31)、新たな設定ファイルを読み込み(ステップS32)、この読み込んだ設定フ  
ァイル内容に応じて再設定を行う(ステップS33)。

【0052】

次に具体例を挙げて説明する。

いま、例えば初期設定ファイルの内容が表1の如きものであったとする。

【0053】

【表1】

	キーボード	マウス
デフォルト	コマンド入力	無 視
文章エディタ	コマンド入力	テキスト入力

【0054】

このような初期設定ファイルを読み込んだ判定部7はデフォルトの場合、キーボード使用時  
にあっては、入力された音声情報をコマンド入力と判定し、またマウス使用時あっては入  
力された音声情報を無視することとなる。

【0055】

また操作者が文章エディタを使用している場合、文章エディタのウィンドウがアクティブ

であれば、キーボード使用時には入力された音声情報をコマンド入力と、またマウス使用時には入力された音声情報をテキスト入力と判定する。

一方このような初期設定ファイルのもとで音声入力装置を使用中に、操作者が波形エディタを使用しようとした場合、この初期設定ファイルで音声波形データの設定が出来ないから設定ファイルの再設定を行う。

いま、再設定のファイルが表2の如くであったとする。

【0056】

【表2】

表 2

	キーボード	マウス
デフォルト	コマンド入力	無 視
文章エディタ	コマンド入力	テキスト入力
波形エディタ	コマンド入力	波 形
波形エディタ（ファイル名入力時）	テキスト入力	テキスト入力

【0057】

これによって、いま波形エディタを使用している状況下では、キーボード使用中の場合には、判定部7は入力された音声情報をコマンド入力と、またマウスを使用中の場合には入力された音声情報を波形入力と夫々判定する。

ただ波形エディタを使用している状況下であっても、ファイル名入力時にはキーボード、マウスのいずれを使用中であっても判定部7は入力された音声情報をテキスト入力と判定することとなる。

【0058】

このような実施例3にあっては判定部7に対し、キーボードの使用中にあっては入力された音声情報を「コマンド」として、またマウス使用中にあっては入力された音声情報を「テキスト」と判定すべく判定のアルゴリズムを設定しておくことで判定部7がこれに従って自動的に判定処理する。これによって操作者の動作と、入力された音声に対する取扱いが協調的となり、作業効率が向上する。

【0059】

【発明の効果】

第1の発明にあっては判定部が音声以外の情報を入力する入力装置が使用中であるか否かに基づいて音声処理部に対してどのような処理を行わせるかを判定することで、この判定結果に基づき音声処理部の処理が自動的に切替えられることとなり、操作者は特別な操作を行うことなく、発声のみで自動処理することが可能となる。

【0060】

第2の発明にあつては判定部が音声以外の情報を入力する入力装置が使用中であるか否か、及び設定ファイルの内容に基づいて音声処理部に対してどのような処理を行わせるかを判定することで、この判定結果に基づき音声処理部の処理が自動的に切替えられることとなり、操作者は特別な操作を行うことなく、発声のみで自動処理することが可能となる。

【0061】

第3の発明にあつては判定部が音声以外の情報を入力する入力装置が使用中であるか否か、及び設定ファイルの内容に基づいて音声処理部に対してどのような処理を行わせるかを判定することで、この判定結果に基づき音声処理部の処理が自動的に切替えられることとなり、操作者は特別な操作を行うことなく、発声のみで自動処理することが可能となる。

【0062】

第4の発明にあつては、音声処理部において音声認識を行う場合には、各種の辞書を操作者が特別な指示を行うことなく、自動的に選定して音声処理部への読出しを可能とする。

【0063】

第5の発明にあつては、判定部が音声以外の情報を入力する入力装置が使用中であるか否かに基づいて音声処理部に対してどのような処理を行わせるかを判定することで、この判定結果に基づき音声処理部の処理が自動的に切替えられることとなり、操作者は特別な操作を行うことなく、発声のみで自動処理することが可能となる。

【0064】

第6の発明にあつては、判定部の判定結果の如何にかかわらず予め定めた特定の音声が入力された場合には、予め定めた最優先順位の処理を行わせることで誤認識が低減される共に、操作性が向上する。

【0065】

第7の発明にあつては、キーボード、マウスを備える汎用コンピュータに広く適用可能となる。

【図面の簡単な説明】

【図1】本発明の原理図である。

【図2】本発明の実施例1の構成を示すブロック図である。

【図3】実施例1における判定部の処理過程を示すフローチャートである。

【図4】実施例2の構成を示すブロック図である。

【図5】実施例2における判定部の処理過程を示すフローチャートである。

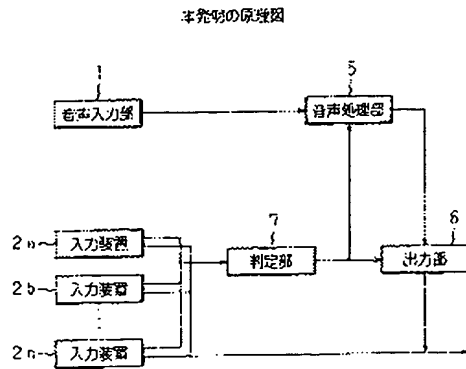
【図6】実施例3における判定部の判定処理内容の初期設定過程及び設定変更過程を示すフローチャートである。

【図7】従来装置の構成を示すブロック図である。

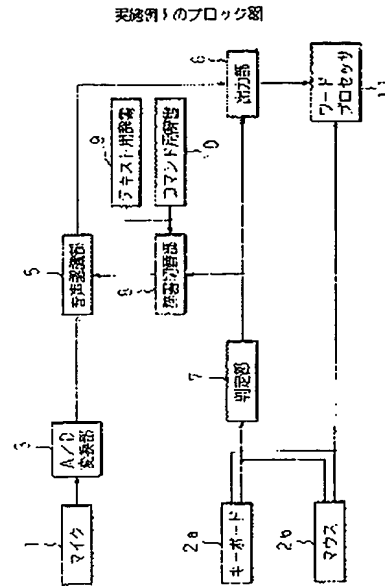
【符号の説明】

- 1 音声入力部
- 2 a～2 n 入力装置
- 5 音声処理部
- 6 出力部
- 7 判定部
- 8 辞書切替部
- 9 テキスト用辞書
- 10 コマンド用辞書
- 11 ワードプロセッサ
- 12 波形エディタ

【図 1】

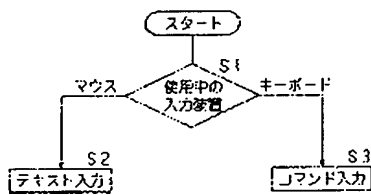


【図 2】



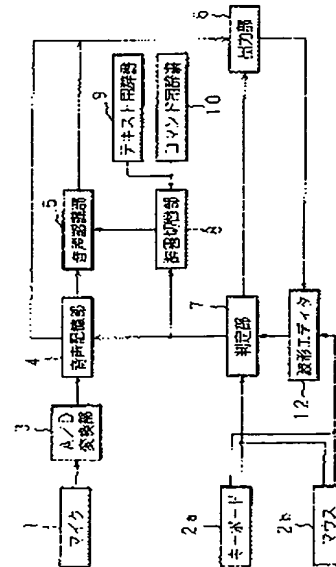
【図 3】

判定部のフローチャート



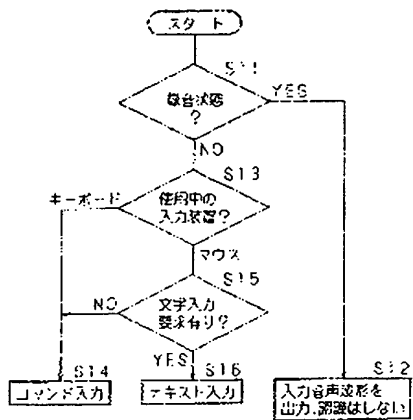
【図 4】

実施例 2 のブロック図



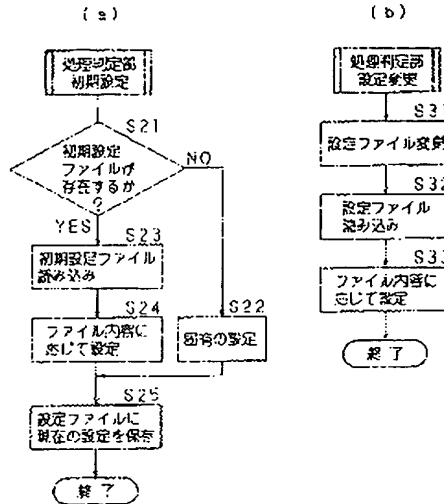
【図 5】

判定部のフローチャート



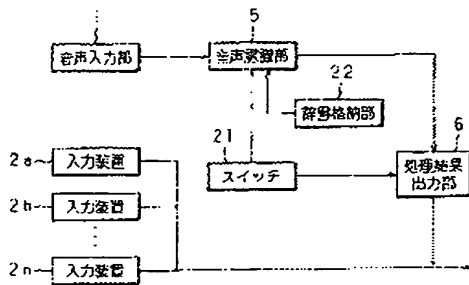
【図 6】

判定部のフローチャート



【図 7】

従来ブロック図



---

フロントページの続き

(58)調査した分野(Int.Cl.<sup>7</sup>, DB名)

G06F 3/16,

3/00, 3/14-3/153,

3/02-3/027,

17/21,

G10L 3/00

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-063330

(43)Date of publication of application : 08.03.1996

(51)Int.Cl. G06F 3/16  
G10L 3/00  
G10L 3/00

(21)Application number : 06-193418

(71)Applicant : FUJITSU LTD

(22)Date of filing : 17.08.1994

(72)Inventor : WASHIO NOBUYUKI

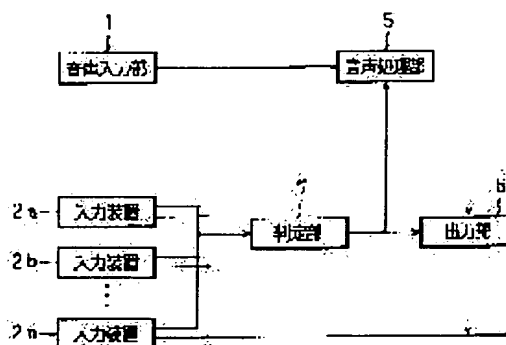
## (54) VOICE INPUT DEVICE

### (57)Abstract:

**PURPOSE:** To attain the automatic switching between the processing contents to be given to the voice information that is inputted next and an output mode based on the using state and/or the using history of an input device which inputs the information other than those on voices.

**CONSTITUTION:** The signal showing a use/non-use state or an under-use state is fetched from each of input devices 2a, 2b...2n such as a keyboard, a mouse, etc., which input the information other than those on voices. Then the fetched signal is sent to a deciding part 7. The part 7 decides the contents of the voice information supplied from a voice input part 1 based on the using states and/or the using histories of the devices 2a, 2b...

2n. Based on the contents of this decision, an instruction is given to a voice processing part 5 to perform a prescribed processing and also an instruction is given to an output part 6 to perform an output operation in a prescribed mode.



## LEGAL STATUS

[Date of request for examination] 03.07.2000

[Date of sending the examiner's decision of

rejection]

[Kind of final disposal of application other than  
the examiner's decision of rejection or  
application converted registration]

[Date of final disposal for application]

[Patent number] 3586777

[Date of registration] 20.08.2004

[Number of appeal against examiner's  
decision of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]



Warning: PAJ Data was not available on download time. You may get bibliographic data in English later.

\* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

## CLAIMS

---

(57) [Claim(s)]

[Claim 1]

In the audio input unit equipped with the voice input section, the input unit which inputs information other than voice, and the speech processing section which performs predetermined processing to the speech information inputted from said voice input section,

The audio input unit which judges the content of the processing which should perform whether said input unit is using it based on the signal of said input unit to the inputted speech information according to the input unit which detected and was detected [ that it is under activity, and ], and is characterized by providing the judgment section which controls said speech processing section according to the judged result.

[Claim 2]

Said judgment section has the configuration file to which the content of the speech information inputted into the input unit detected [ that it is under activity and ] was set,

The audio input unit according to claim 1 which judges the content of the processing which should be performed to the inputted speech information based on the input unit detected [ that it is under activity, and ], and said configuration file, and is characterized by controlling said speech processing section according to the judged result.

[Claim 3]

Said configuration file matches speech information with application a default or in use, and is set up,

Said judgment section detects the signal of application in use, and it judges whether the existence of application in use and application are using it,

JP-B-3380771 2/20

The audio input unit according to claim 2 which judges the content of the processing which should be performed to the inputted speech information based on the input unit detected [ that it is under activity, and ], the application judged that are under activity, and said configuration file, and is characterized by controlling said speech processing section according to the judged result.

[Claim 4]

Said speech processing section is equipped with two or more voice dictionaries and the dictionary change section which chooses any one or the plurality of two or more of these voice dictionaries,

It is an audio input unit given in claim 1 which said judgment section outputs the selection command of a voice dictionary to said dictionary change section according to the judged result, and is characterized by said dictionary change section changing a voice dictionary based on said selection command thru/or any 1 term of 3.

[Claim 5]

In the audio input unit equipped with the voice input section, the input unit which inputs information other than voice, the speech processing section which performs predetermined processing to the speech information inputted from said voice input section, and the output section which outputs the result processed in this speech processing section,

The audio input unit characterized by to provide the judgment section which judges the output mode of the content of processing which should perform whether said input unit is using it based on the signal of said input unit to the inputted speech information according to the input unit which detected and was detected [ that it is under activity, and ], and this processing result, and controls said speech-processing section and said output section according to this judgment result.

[Claim 6]

It is the audio input unit according to claim 5 characterized by making it have operated in order to output the mode which defined said output section beforehand regardless of the judgment result of the judgment section when the recognition result of the speech processing section to the inputted speech information was the word defined beforehand.

[Claim 7]

Said input device is an audio input unit given in claim 1 characterized by being a keyboard and/or a mouse thru/or any 1 term of 6.

---

[Translation done.]

JP-B-5580777

07/20

\* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Industrial Application]

This invention relates to the audio input unit which enabled the change of modification of the content of an output of the speech information into which the content of processing which should be given to the inputted speech information was changed and inputted automatically, without using a switch etc.

[0002]

[Description of the Prior Art]

Drawing 7 is the block diagram showing the configuration of the audio input unit in the former, and one in drawing is the voice input sections, such as a microphone, 2a, and 2b. -- 2n of input units which input other information other than voice, such as a keyboard and a mouse, is shown.

The speech information inputted from the voice input section 1 is inputted into the speech recognition section 5.

The speech recognition section 5 is set as the processing mode according to each, such as the speech information beforehand inputted with a switch 21, for example, text information, and command information. When a processing mode is text information processing Mohd, the dictionary for text information processing is read from the dictionary storing section 22. Based on this, in being command information processing Mohd, it reads the dictionary for command information processing from the dictionary storing section 22, and recognition processing of the speech information inputted based on this is performed, and a recognition result is outputted to the processing result output section 6.

The processing result output section 6 is also set as the output mode

corresponding to the speech information beforehand inputted with a switch 21, and outputs the inputted recognition result as a command as a text with the input from other input units 2a-2n, respectively.

[0003]

[Problem(s) to be Solved by the Invention]

By the way, therefore, the object inputted through the voice input section 1 changes each time, when it is text, such as a text, and when it is the operating instructions to application, a window manager, and OS, or when it is a voice data point.

[0004]

Since the content of the processing which should be performed in the speech recognition section 5 to such various objects for an input, and the procedure of processing also differed from self-\*\*, if the speech recognition section 5 needed to be switched to the processing mode which was adapted for each object for an input and it was in the former, the switch 21 was switched by hand control or voice input, and the processing mode was set up. This point is the same also in the processing result output section 6.

[0005]

A user a switch 21 to switch with hand control However, a keyboard in use, or -- from a mouse etc. -- once -- a hand -- not detaching, in order not to obtain, to interrupt actuation of a keyboard and a mouse and to switch by voice input When it was naturally necessary to register the special command for a change, the incorrect recognition resulting from the conversation in the circumferences other than a noise and other input voice etc. arose, and when the operator did not expect, there was inconvenience -- the change of a processing mode and output mode may be performed suddenly.

[0006]

This invention is made in view of this situation, and the place made into the object is to enable it to make automatically modification of the processing which should be performed to input voice, and a change of an output mode, without requiring actuation special to an operator.

Other objects of this invention are to enable it to switch a voice dictionary automatically according to the speech information which performed speech recognition processing to the speech information inputted in the speech processing section, and was inputted

[ command / the text, ] into the case.

[0007]

The object of further others of this invention is to judge this

JP-B-5500777 5/20  
automatically in the judgment section and enable control to the output section, also when outputting to the inputted speech information, without performing speech processing.

Although the usual computers, such as a keyboard and a mouse, are equipped with the object of further others of this invention, it is to enable far-reaching application by supposing that the judgment section judges based on an operating condition and/or activity hysteresis.

[0008]

When the inputted speech information is the word defined beforehand, regardless of the judgment result of the judgment section, the object of further others of this invention is making the mode defined beforehand output from the output section, and is shown in attaining the increase in efficiency of processing.

The object of further others of this invention is to enable modification of the content of the judgment processing in the judgment section if needed, and is to close applicability, if amplification is still more possible.

[0009]

[Means for Solving the Problem]

It explains based on principle drawing showing the principle of this invention in drawing 1 below.

Drawing 1 is principle drawing of the audio input unit concerning this invention, and one in drawing shows input units which input information other than voice, such as a keyboard and a mouse, the voice input section, 2a, and 2b-2n.

The speech information inputted from the voice input section 1 is inputted into the speech processing section 5 as digital information.

[0010]

On the other hand, input units [ 2a-2n ] an operating condition and/or activity hysteresis are crowded judgment section 7 HE picking serially.

The content of the judgment processing to which the judgment section 7 was beforehand set based on these operating conditions and/or activity hysteresis, The speech information inputted into the speech information currently actually inputted from the voice input section 1 according to an algorithm or a degree namely, in the thing of what kind of content for example, text input or command input -- or data without the need of processing in any way in the speech processing section etc. are judged, and a predetermined command is given to the speech processing section 5 and the output section 6 based on this judgment result, respectively.

[0011]

Generally, in the case of the operator who became skillful in the input of a keyboard, usually, the direction inputted using a keyboard can process promptly rather than inputting a text etc. by voice input. Therefore, when using the keyboard, objects it is considered that want to dare carry out voice input in the skilled operator are operating instructions, such as opening of a window, and a mode change of application, in many cases. On the other hand, when a mouse is under activity, to inputting a text etc., a hand is once lifted from a mouse, and a text is inputted using a keyboard, and since the actuation which returns a hand to a mouse is needed again, if you want to input text not only by operating instructions but by voice input, it will want in many cases.

That is, it becomes possible to judge the output mode of the content of processing, and the output section 6 which should be given in the speech processing section 5 to input voice by grasping the keyboard which is the input device which inputs information other than voice, the activity hysteresis of a mouse, and an operating condition.

[0012]

To the speech processing section 5, to the speech information inputted, when speech recognition processing should be performed [ whether speech recognition processing should be performed and ] again, as a text, the command of whether to perform recognition processing as a command is given, and the speech processing section 5 is controlled. Moreover, that it should output as a command that it should output as a text to the output section 6 when the recognition result given from the speech processing section 5 is a text when it is a command, about the content from which it was not further processed at all in the speech processing section 5, a command is given that this should be outputted to a wave editor, for example, and the output section 6 is controlled.

[0013]

The speech processing section 5 outputs this to the output section 6 by this, without performing predetermined speech processing to the speech information which mode setting of the response of was made possible to the speech information inputted according to the command from the judgment section 7, and was inputted, or processing.

Moreover, similarly the output section 6 follows a command from the judgment section 7, and, in the case of text input, the input from the speech processing section 5 is input-device 2a of the others [ case / of the command input as a text ] as command input, and 2b. -- It outputs to a word processor, a wave editor, etc. like the input from 2n.

[0014]

JP-B-3500777 7/20

In addition, when the recognition result of the inputted speech information is the specific word defined beforehand, the output section is good also as giving priority over other modes to the output of the mode defined beforehand, and performing it.

Moreover, an operator's content of judgment processing of the judgment section 7 is possible for modification to arbitration, and an operator enables it to have expanded the applicable range by making a judgment result change according to the object for application.

[0015]

In the audio input unit with which the audio input unit concerning the 1st invention was equipped with the speech processing section which performs predetermined processing to the voice input section, the input unit into which information other than voice is inputted, and the speech information into which it was inputted from said voice input section The content of the processing which should perform whether said input unit is using it based on the signal of said input unit to the inputted speech information according to the input unit which detected and was detected [ that it is under activity and ] is judged, and it is characterized by providing the judgment section which controls said speech processing section according to the judged result.

[0016]

The audio input unit concerning the 2nd invention is characterized by to have the configuration file to which the content of the speech information inputted into the input unit detected [ that said judgment section is under activity in the 1st invention and ] was set, to judge the content of the processing which should perform to the inputted speech information based on the input unit detected [ that it is under activity, and ], and said configuration file, and to control said speech-processing section according to the result which judged.

[0017]

The audio input unit concerning the 3rd invention is set to the 2nd invention. Said configuration file Speech information is matched with application a default or in use, and it has set up. Said judgment section Detect the signal of application in use and it judges whether the existence of application in use and application are using it. Based on the input unit detected [ that it is under activity and ], the application judged that are under activity, and said configuration file, the content of the processing which should be performed to the inputted speech information is judged, and it is characterized by controlling said speech processing section according to the judged result.

[0018]

The audio input unit which the 4th invention requires is set to the 1st thru/or the 3rd invention. The voice dictionary of plurality [ section / said / speech processing ], It has the dictionary change section which chooses any one or the plurality of two or more of these voice dictionaries, said judgment section outputs the selection command of a voice dictionary to said dictionary change section according to the judged result, and said dictionary change section is characterized by changing a voice dictionary based on said selection command.

[0019]

The input unit into which the audio input unit concerning the 5th invention inputs information other than voice as the voice input section, In the audio input unit equipped with the speech processing section which performs predetermined processing to the speech information inputted from said voice input section, and the output section which outputs the result processed in this speech processing section It detects whether based on the signal of said input unit, said input unit is using it. According to the input unit detected [ that it is under activity and ], the output mode of the content of processing which should be given to the inputted speech information, and this processing result is judged, and it is characterized by providing the judgment section which controls said speech processing section and said output section according to this judgment result.

[0020]

When the audio input unit concerning the 6th invention is the word which the recognition result of the speech processing section to the inputted speech information defined beforehand in the 5th invention, said output section is characterized by making it have operated in order to output the mode defined beforehand regardless of the judgment result of the judgment section.

[0021]

The audio input unit concerning the 7th invention is characterized by said input devices being a keyboard and/or a mouse in the 1st thru/or the 6th invention.

[0022]

[Function]

If it is in the 1st invention, the processing which the judgment section should perform to input voice according to whether the input unit which inputs information other than voice is using it will be judged automatically, it will direct to the speech processing section, and an



operator does not need directions of the content of processing.

[0023]

If it is in the 2nd invention, the processing which the judgment section should perform to input voice based on the content of whether the input unit which inputs information other than voice is using it, and the configuration file will be judged automatically, it will direct to the speech processing section, and an operator does not need directions of the content of processing.

[0024]

If it is in the 3rd invention, the processing which the judgment section should perform to input voice based on the content of whether the input unit which inputs information other than voice is using it, and the configuration file will be judged automatically, it will direct to the speech processing section, and an operator does not need directions of the content of processing.

[0025]

If it is in the 4th invention, a change becomes possible automatically in the dictionary change section about two or more dictionaries used in the speech processing section for speech recognition.

[0026]

If it is in the 5th invention, it becomes possible to perform control to these automatically with the judgment section outputting a command to the speech processing section and each output section according to a judgment result based on whether the input unit which inputs information other than voice is using it.

[0027]

If it is in the 6th invention, it is making the mode beforehand set to the output section output to the recognition result beforehand defined regardless of the judgment result of the judgment section, and operability may be improved while reducing malfunction.

[0028]

If it is in the 7th invention, it is using the information about whether a keyboard and a mouse using it, and becomes applicable to a general purpose computer equipped with a keyboard and a mouse.

[0029]

[Example]

(Example 1)

This invention is concretely explained based on the drawing in which the example is shown below.

Drawing 2 is the block diagram showing the configuration at the time of

applying the audio input unit concerning this invention to the word processor 11 with a graphic form edit function, and the voice input section by which one in drawing was constituted from a microphone etc., 2a, and 2b show input units which input information other than voice, such as a keyboard and a mouse.

The speech information inputted from the voice input section 1 is inputted into the speech processing section 5 which the analog signal was changed by the digital signal and consisted of the A/D-conversion sections 3 as the speech recognition section.

[0030]

Input unit 2a which, on the other hand, inputs information other than voice, and the input from 2b are inputted into WORD PUSSESSA 11, and also they are incorporated serially to the judgment section 7.

The judgment section 7 judges the content of the processing which should be performed to the content of the speech information which recognizes these operating conditions and/or activity hysteresis according to input unit 2a, such as a keyboard and a mouse, and the content of judgment processing beforehand set up based on the signal from 2b, i.e., an algorithm, and is actually inputted through the voice input section 1, and is inputted behind, and the speech information inputted.

The output mode from the content and the output section 6 of the processing which the specifically inputted speech information should perform to text information, command information, and a list in the speech processing section 5 to each information is judged, and the command according to each is given to the dictionary change section 8 and the output section 6.

[0031]

In addition, when the sensor which detects the existence of that activity is attached to each input unit 2a, such as a keyboard and a mouse, and 2b, it is good also as incorporating this sensor output in the judgment section 7, and judging by recognizing these operating conditions and activity hysteresis.

The dictionary change section 8 reads the dictionary for texts, or the dictionary 10 for commands to the speech processing section 5 by the command from the judgment section 7.

[0032]

The speech processing section 5 reads the dictionary 9 for texts or the dictionary 10 for /commands selectively changed in the dictionary change section 8 which operates based on the command from said

judgment section 7, performs recognition processing of speech information based on these, and outputs a recognition result to the output section 6. Based on the directions from said judgment section 7, in the case of the command input as a text, the output section 6 outputs this to a word processor 11 as a command, when speech information is text input.

[0033]

Next, it explains according to the flow chart which shows actuation of this invention equipment to drawing 3.

Drawing 3 is a flow chart which shows a processing process in case the judgment section 7 judges based on what an input device present in use is. first -- input unit 2a in use and 2b -- or [ any ] -- judging -- (step S1) text input [ as opposed to / in the speech information inputted when an input unit in use is a mouse / the document under edit with a word processor 11 ] -- judging -- (step S2) Moreover, in being a keyboard, the speech information inputted judges with the command input to a word processor 11. (step S3) The directions which correspond to the dictionary change section 8 and the output section 6 based on each judgment are outputted.

[0034]

Next, an example is given and the content of processing is explained. For example, if it utters with "save" in order to take backup of a text when an operator is creating a text with a word processor 11 using a keyboard Recognize that an operator is keyboard using the judgment section 7, and the inputted speech information judges it like the above-mentioned to be the command input to a word processor 11. A command is outputted that the dictionary 10 for commands should be chosen to the dictionary change section 8, and it directs by making the recognition result of the speech recognition section into a command to the output section 6 that it should output to a word processor 11.

[0035]

Consequently, the speech information inputted in the speech processing section 5 constituted as the speech recognition section is recognized to be "save" using the dictionary 10 for commands, and that recognition result is outputted to the output section 6. The output section 6 makes a recognition result "save" a command "save", it outputs to a word processor 11, and a word processor 11 saves the document under edit in response to a command "save."

[0036]

Moreover, a location to put "Japanese" first is specified with a mouse in

order to write the text "Japanese" in the predetermined part in the graphic form as that to which the operator is performing graphic form edit with the word processor 11, and suppose that it uttered with "Japanese." The judgment section 7 judges the voice inputted as the operator recognizes and mentioned that a mouse was using it above to be the text input to the document under edit of a word processor 11, and directs it that the dictionary 9 for texts should be chosen as the dictionary change section 8, and is directed that it should output as a text display to the output section 6.

[0037]

By this, the speech processing section 5 recognizes the inputted speech information to be "Japanese" using the dictionary 9 for texts, and outputs this recognition result to the output section 6. The output section 6 uses "Japanese" as a text, it outputs to a word processor 11 and a word processor 11 indicates the "Japanese" which is a text by insertion in the directions location by the mouse.

[0038]

(Example 2)

The example 2 is performing edit to the speech information inputted using the wave editor 12, and shows the case where there is no need that the speech processing section 5 constituted as the speech recognition section to the inputted speech information performs special processing.

Drawing 4 is the block diagram showing the configuration of the example 2 of this invention. In this example 2, the phonetic memory section 4 is infixed in the medium of the A/D-conversion section 3 and the speech processing section 5 constituted as the speech recognition section, and it is made to have outputted the directions from the judgment section 7 also to this phonetic memory section 4 besides the dictionary change section 8 and the output section 6. Moreover, it is made to be given to the judgment section 7 in the signal which shows that the output from the output section 6 other than the output from input-device 2a, such as a keyboard and a mouse, and each 2b is inputted, and the wave editor 12 is [ wave editor ] under activity from the wave editor 12.

[0039]

The judgment section 7 outputs a command that it should recognize that it is data for wave edit of the content of the inputted speech information based on input-device 2a, such as a keyboard and a mouse, the signal which shows the activity of 2b, and the signal which shows the activity of the wave editor from the wave editor 12, and a command should

output that voice should be memorized to the phonetic-memory section 4, and the wave should output to the wave editor 12 to the output section 6.

Drawing 5 is a flow chart which shows the processing process of the judgment section 7. first -- the case where the speech information judge that is the object which does not perform recognition processing of as opposed to this only by outputting the speech information (wave) inputted when the inputted speech information judged whether it was under [ sound-recording ] \*\*\*\*\* in the phonetic-memory section 4 (step S11) and was recording (step S12) and which was inputted on the other hand is not being recorded -- an input device in use -- a keyboard -- or a mouse judges (step S13).

[0040]

When the inputted speech information is judged to be command input in the case of a keyboard (step S14) and a mouse is under activity, it judges whether there is any demand of an alphabetic character input (step S15), when there is nothing, the inputted speech information is judged to be command input (step S14), and in being, the inputted speech information judges with text input (step S16).

[0041]

An operator inputs his voice through a microphone concretely (sound recording), the wave is edited, and the case where it is under activity that an edit result should be saved at a file is explained.

An operator utters toward a microphone and starts the sound recording of his voice. Under such a condition, the information on the purport which voice is recording from the wave editor 12 to the judgment section 7 is inputted. By this, to the speech information inputted in the speech processing section 5, the judgment section 7 does not need to perform audio recognition processing, and judges only with it being that the wave may only be outputted.

The judgment section 7 is directed that the voice wave which ordered that the speech information inputted to the phonetic memory section 4 based on this judgment should be recorded, and was inputted to the output section 6 should be outputted to the wave editor 12 as it is. In addition, since speech recognition processing is not needed to the dictionary change section 8, a command is not outputted.

[0042]

Consequently, after the speech information digitized in the A/D-conversion section 3 is recorded in the phonetic memory section 4, it is outputted to the direct-output section 6, and the output section 6

outputs the inputted voice wave to the wave editor 12.

An operator starts wave-like edit, after the sound recording of utterance is completed. The wave editor 12 processes that the wave specified supposing it performed echo processing to the wave which the operator specified the wave-like section using the mouse, and was specified that it utters with an "echo" and uttered with "the clearance" should be eliminated.

[0043]

That is, when a current keyboard is using it, the judgment section 7 judges with the speech information inputted based on the signal inputted through the signal and the wave editor 12 in which it is shown that it is under activity from a keyboard being a command.

It orders in order, as for the judgment section 7, to send voice to the speech processing section 5 to the phonetic memory section 4 according to this, and it orders that the dictionary 10 for commands should be chosen to the dictionary change section 8, and it orders in order to send a command to the wave editor 12 to the output section 6.

[0044]

Consequently, the speech processing section 5 performs recognition processing to the speech information inputted using the dictionary 10 for commands, recognizes the "echo" or "a clearance" which is input voice, and outputs this to the output section 6.

The output section 6 makes a command the "echo" or "a clearance" which it is as a result of recognition, and delivery and this command are executed to the wave editor 12.

Next, suppose that it uttered with "save" that the content which the operator edited should be saved. If this "save" name-unestablished-law-files and puts in another way, when it will be a new file, the wave editor 12 requires the name of a file. Then, when it utters with "its voice" as a file name, even if a mouse is under activity, since the wave editor 12 requires text input, it judges the speech information into which the judgment section 7 was inputted to be a text.

[0045]

The judgment section 7 is directed that a command should be outputted that the dictionary 9 for texts should be chosen to the dictionary change section 8, and "its voice" as a text should be outputted to the output section 6. Consequently, the speech processing section 5 performs recognition processing to speech information using the dictionary 9 for texts, and outputs this to the output section 6. The output section 6 is outputted to the wave editor 12 by using as a text "one's voice" which

it is as a result of recognition, and "its voice" which is a file name is inputted into the wave editor 12, and it is saved.

If it is in such an example 2, even if the alphabetic character input like a file name or the numerical input of the delay time like an "echo" does not move a hand from the mouse under actuation to a keyboard, the input of it is attained by voice input.

[0046]

In addition, no matter the judgment result of the judgment section 7 may be what thing, when the result of speech recognition is "the specific word" defined beforehand, the output section 6 is good about which [ of examples 1 and 2 ] case, also as performing the predetermined output control defined beforehand.

For example, when a specific word is "resizing" which is the operator command to a window manager and OS, or a predetermined name of a person, for example, "Tanaka", if there is the output section 6 in "resizing", it outputs the operator command for size change of a window to a window manager and OS.

[0047]

when the speech information into which application was under actuation under the predetermined window, and the judgment section 7 was inputted is judged to be a command to application in "resizing", it becomes equivalent to having set up the priority of the content of processing over input voice substantially using a recognition result, and operability is markedly alike and improves.

Moreover, the misconception depended also when there is a possibility that the voice of "Tanaka" may mix frequently from a perimeter by setting up the content of processing of the judgment section so that [ "so that this may be disregarded" ], when "Tanaka" who is the name of a person currently used widely is inputted from the voice input section 1 can be avoided.

[0048]

(Example 3)

Although the content of judgment processing of what kind of judgment to make perform based on input unit 2a, the operating condition of 2b, and activity hysteresis to the judgment section 7 in the examples 1 and 2 explained the case where it was initialized on the occasion of start up of a voice input system, it can be reset in this example 3 at the event of arbitration.

[0049]

The content [ in / in drawing 6 (a) / the judgment section 7 ] of

judgment processing, i.e., the initialization processing process of an algorithm, and drawing 6 (b) are the flow charts of each setting-out modification processing process of an algorithm.

first, a \*\*\*\*\* [ that, as for initial setting of an algorithm, an initialization file exists at the time of starting of an audio input unit ] -- judging -- (step S21) the case where it does not exist -- the time of "setting out of a proper", for example, a keyboard activity, -- command input -- moreover, judgment processing made into text input is set up at the time of a mouse activity (step S22) .

[0050]

Moreover, when the configuration file which a user will set up according to liking if it puts in another way when it exists exists, priority is given over "setting out of a proper" in the first half, and the judgment section 7 reads an initialization file. (step S23) It sets up according to the content of this initialization file. (step S24) The current content of setting out is saved at a configuration file. (step S25) .

[0051]

On the other hand, in resetting, a user changes a configuration file using an editor etc. (step S31), and a new configuration file is reset according to read in (step S32) and this read content of a configuration file (step S33).

[0052]

Next, an example is given and explained.

Suppose that the content of the now, for example, initialization, file was a thing like a table 1.

[0053]

[A table 1]

表 1

	キーボード	マウス
デフォルト	コマンド入力	無 視
文章エディタ	コマンド入力	テキスト入力



[0054]

The speech information inputted if the judgment section 7 which read such an initialization file judged the speech information inputted if it was at the time of a keyboard activity in the default case to be command input and it was at the time of a mouse activity will be disregarded.

[0055]

Moreover, if the window of a text editor is active when the operator is using the text editor, at the time of a keyboard activity, the speech information into which the inputted speech information was inputted with command input again at the time of a mouse activity will be judged to be text input.

On the other hand, under such an initialization file, when an operator tries to use a wave editor while using the audio input unit, since setting out of a voice data point cannot be performed in this initialization file, a configuration file is reset.

Now, suppose that the file of resetting was as shown in a table 2.

[0056]

[A table 2]

表 2

	キーボード	マウス
デフォルト	コマンド入力	無 視
文章エディタ	コマンド入力	テキスト入力
波形エディタ	コマンド入力	波 形
波形エディタ（ファイル名入力時）	テキスト入力	テキスト入力

[0057]

JP-B-3500777 10/20

By this, the speech information inputted in the speech information into which the judgment section 7 was inputted when it was [ keyboard ] under activity when a mouse was under activity with command input again is judged under the situation of using the wave editor now, respectively to be a wave input.

Even if it is under the situation of merely using the wave editor, and it is using any of a keyboard and a mouse at the time of a file name input, the judgment section 7 will judge the inputted speech information to be text input.

[0058]

The judgment section 7 carries out judgment processing automatically according to this by setting up the algorithm of a judgment that the speech information inputted if it was during the mouse activity should be judged to be a "text" by making into a "command" speech information inputted if it was in such an example 3 and was to the judgment section 7 while using a keyboard. By this, the handling to the voice inputted as actuation of an operator becomes cooperative, and working efficiency improves.

[0059]

[Effect of the Invention]

If it is in the 1st invention, it is judging what kind of processing being made to perform to the speech processing section based on whether the input unit into which the judgment section inputs information other than voice is using it, based on this judgment result, processing of the speech processing section will be changed automatically, and an operator becomes possible [ carrying out automatic processing only by utterance ], without performing special actuation.

[0060]

If it is in the 2nd invention, it is judging what kind of processing being made to perform to the speech processing section based on the content of whether the input unit into which the judgment section inputs information other than voice is using it, and the configuration file, based on this judgment result, processing of the speech processing section will be changed automatically, and an operator becomes possible [ carrying out automatic processing only by utterance ], without performing special actuation.

[0061]

If it is in the 3rd invention, it is judging what kind of processing being made to perform to the speech processing section based on the content of whether the input unit into which the judgment section inputs

information other than voice is using it, and the configuration file, based on this judgment result, processing of the speech processing section will be changed automatically, and an operator becomes possible [ carrying out automatic processing only by utterance ], without performing special actuation.

[0062]

If it is in the 4th invention, in performing speech recognition in the speech processing section, various kinds of dictionaries are selected automatically, without an operator performing special directions, and it makes possible read-out to the speech processing section.

[0063]

It is judging what kind of processing being made to perform to the speech processing section based on whether if it is in the 5th invention, the input unit into which the judgment section inputs information other than voice is using it, based on this judgment result, processing of the speech processing section will be changed automatically, and an operator becomes possible [ carrying out automatic processing only by utterance ], without performing special actuation.

[0064]

if it is in the 6th invention, when the specific voice beforehand defined regardless of the judgment result of the judgment section is inputted, incorrect recognition is reduced by making the ranking of the highest priority defined beforehand process -- operability both improves.

[0065]

If it is in the 7th invention, it becomes it can be large to a keyboard and a general purpose computer equipped with a mouse, and applicable to them.

[Brief Description of the Drawings]

[Drawing 1] It is principle drawing of this invention.

[Drawing 2] It is the block diagram showing the configuration of the example 1 of this invention.

[Drawing 3] It is the flow chart which shows the processing process of the judgment section in an example 1.

[Drawing 4] It is the block diagram showing the configuration of an example 2.

[Drawing 5] It is the flow chart which shows the processing process of the judgment section in an example 2.

[Drawing 6] It is the flow chart which shows the initialization process and setting-out modification process of the content of judgment processing of the judgment section in an example 3.

[Drawing 7] It is the block diagram showing the configuration of equipment conventionally.

[Description of Notations]

1 Voice Input Section

2a-2n Input unit

5 Speech Processing Section

6 Output Section

7 Judgment Section

8 Dictionary Change Section

9 Dictionary for Texts

10 Dictionary for Commands

11 Word Processor

12 Wave Editor

---

[Translation done.]

\* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is principle drawing of this invention.

[Drawing 2] It is the block diagram showing the configuration of the example 1 of this invention.

[Drawing 3] It is the flow chart which shows the processing process of the judgment section in an example 1.

[Drawing 4] It is the block diagram showing the configuration of an example 2.

[Drawing 5] It is the flow chart which shows the processing process of the judgment section in an example 2.

[Drawing 6] It is the flow chart which shows the initialization process and setting-out modification process of the content of judgment

processing of the judgment section in an example 3.

[Drawing 7] It is the block diagram showing the configuration of equipment conventionally.

[Description of Notations]

1 Voice Input Section

2a-2n Input unit

5 Speech Processing Section

6 Output Section

7 Judgment Section

8 Dictionary Change Section

9 Dictionary for Texts

10 Dictionary for Commands

11 Word Processor

12 Wave Editor

---

[Translation done.]

\* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

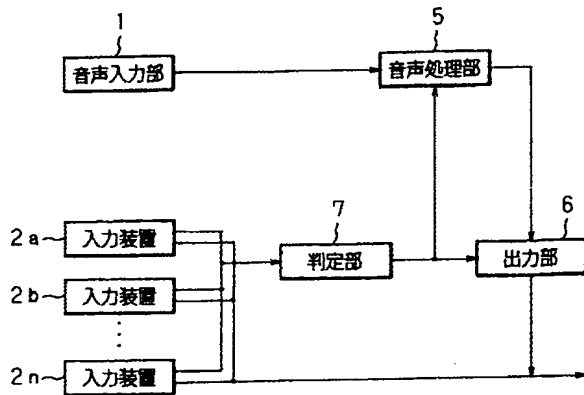
---

DRAWINGS

---

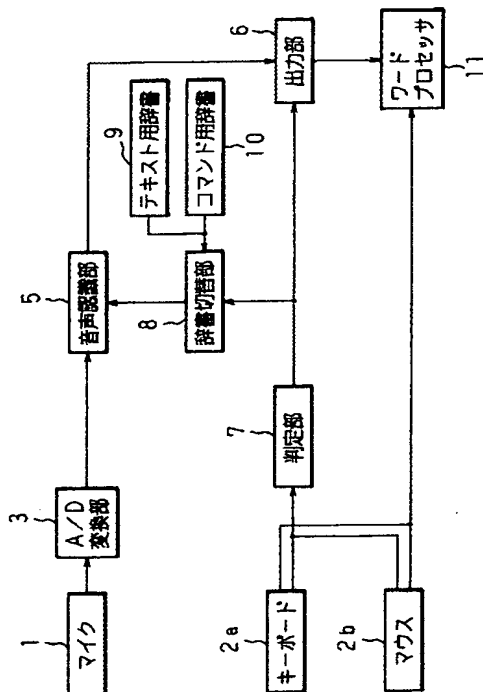
[Drawing 1]

本発明の原理図



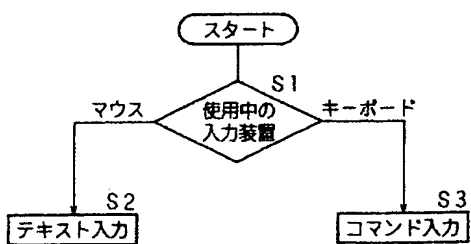
[Drawing 2]

実施例1のブロック図



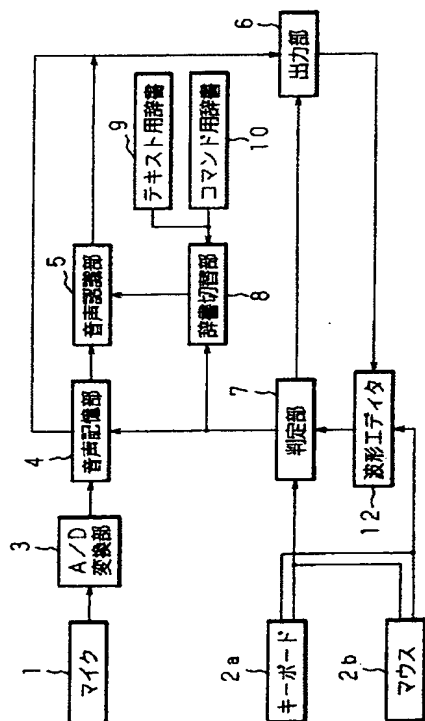
[Drawing 3]

判定部のフローチャート



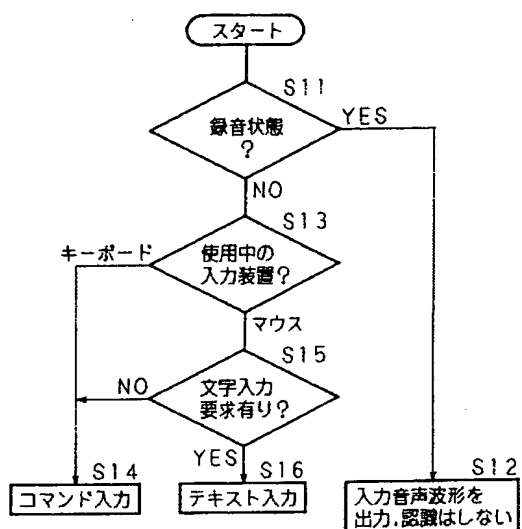
[Drawing 4]

実施例2のブロック図



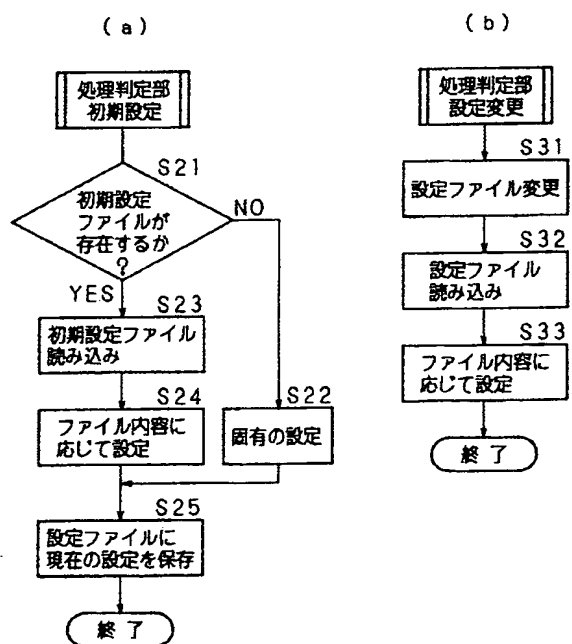
[Drawing 5]

判定部のフローチャート



[Drawing 6]

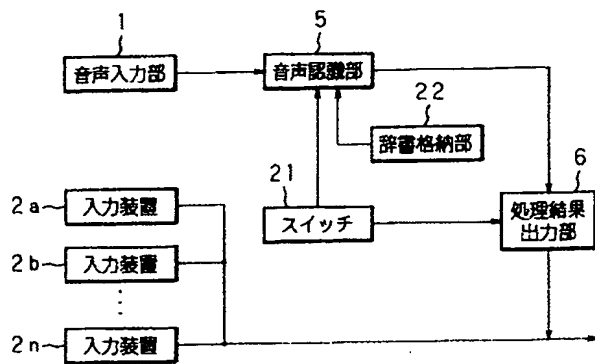
判定部のフローチャート



[Drawing 7]



従来のブロック図



[Translation done.]